IOP Conf. Series: Journal of Physics: Conf. Series **1037** (2018) 032024 doi:10.1088/1742-6596/1037/3/032024

Wind turbine $C_{\mathrm{p,max}}$ and drivetrain-losses estimation using Gaussian process machine learning

E Hart[1], W E Leithead[2], J Feuchtwang[3]
[1] Wind and Marine Energy Systems CDT, University of Strathclyde, Glasgow, UK
[2] Centre for Industrial Control, University of Strathclyde, Glasgow, UK
[3] Institute for Energy and the Environment, University of Strathclyde, Glasgow, UK

E-mail: edward.hart@strath.ac.uk

Abstract. In this paper it is shown that measured data in a wind turbine, available to the controller, can be formulated into a polynomial regression problem in order to estimate the turbine's maximum efficiency power coefficient, $C_{\mathrm{p,max}}$, and drivetrain losses, assuming the latter can be well approximated as being linear. Gaussian process (GP) machine learning is used for the regression problem. These formulations are tested on data generated using the Supergen Exemplar 5 MW wind turbine model, with results indicating that this is a potential low cost method for detecting changes in aerodynamic efficiency and drivetrain losses. The GP approach is benchmarked against standard least-squares (LS) regression, with the GP shown to be the superior method in this case.

1. Introduction

As wind turbine assets grow in size and move further offshore, where access becomes more problematic and expensive, the need for increased reliability becomes even more pronounced. Therefore, there is a need for a diverse range of monitoring techniques with which faults and behavioural changes in wind turbines can be quickly detected and appropriate steps taken. Due to this, there have been a huge number of new measurement techniques developed in recent years, all with their associated costs. In this paper, rather than proposing a new measurement technique per se, we instead look at how existing data available to a wind turbine controller can be more fully utilised in order to try and detect changes in a turbine's aerodynamic efficiency and drivetrain losses.

Changes in aerodynamic efficiency can either happen rapidly, for example via icing of the blades [1], or slowly, for example via blade degradation [2]. Changes in the mechanical losses in the drivetrain can be indicative of damage or an imminent failure. In all cases, prior warning of any changes, and the tracking of such changes over time, is desireable to inform operation and maintenance (O&M) for these wind turbines.

The proposed method uses Gaussian process machine learning to perform polynomial regression on noisy data available to the wind turbine controller. As will be shown, this noisy polynomial data will have coefficients from which estimates of aerodynamic efficiency and drivetrain losses can be formed. This technique does not require any new measurements or equipment, but only the utilisation of existing processing power within the turbine or control room. As such, it presents the possibility for an incredibly low cost addition to the monitoring toolbox of any wind turbine operator.

Gaussian process machine learning has already seen a range of applications in wind energy, including forecasting [8], structural health monitoring [9], and nonlinear dynamics identification [10].

2. Control Regions

The control strategy for a wind turbine generally contains two distinct regions while operational. Let 'rated wind speed' be the wind speed at which the wind turbine reaches its rated power output. Then the two major control regions are 'below rated' and 'above rated' operation. In this work we consider variable speed, variable pitch (VSVP) machines which represent the vast majority of the currently operational turbines globally. For such a turbine, a typical design power curve is given in Figure 1.

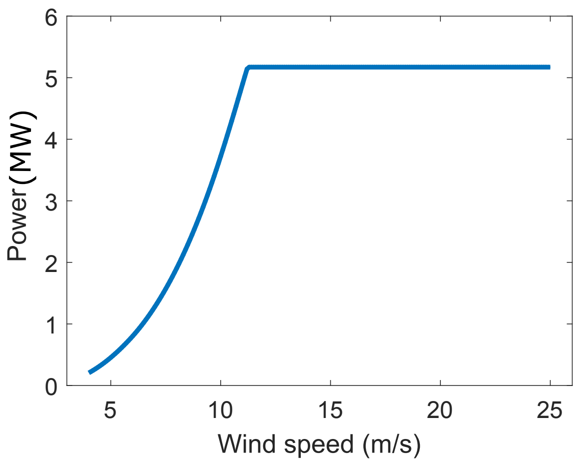

Figure 1. Typical design power curve for a VSVP wind turbine. The rated power here is 5 MW. Note, this is the power curve for the Supergen turbine model used to generate data for the current work.

The rated wind speed can be seen to be roughly 11.5 m/s. Above this point the turbine will be pitching in order to limit the power, while maintaining a constant rotational speed. This work, however, is focussed on the below rated portion of the power curve; where the turbine is attempting to maintain optimum efficiency in order to maximise power capture. This corresponds to maintaining the tip-speed ratio, λ, defined as,

$$\lambda = \frac{\omega_r R}{v}, \tag{1}$$

at the value which corresponds to maximum efficiency, this value will be denoted by $\lambda_{\max}$. From Equation 1 it should be clear that maintaining $\lambda = \lambda_{\max}$ requires the turbine rotational speed to change as the wind speed changes. The design strategy for a wind turbine is commonly represented on a torque-speed diagram [3]. The reason for using these diagrams is that they are useful design tools since torque control is used to vary rotational speed in below rated operation and hence both these parameters are represented. Futhermore, structural modes and generator design limits can also be plotted on the torque-speed diagram, making it easy to design a strategy which takes these constraints into account. One such diagram is shown in Figure 2 for the Supergen turbine model used in the current work. In this paper we are interested in the variable speed section at the centre of the diagram which can be seen to track the maximum efficiency curve. Note also that Figure 2 contains a sample trajectory from a simulation of the given turbine model. For a detailed explanation of how the turbine tracks the maximum efficiency curve see, for example, [4].

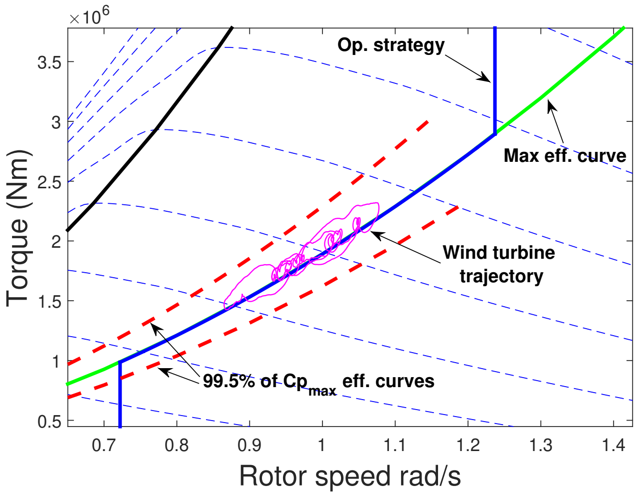

Figure 2. Wind turbine torque-speed diagram including maximum efficiency curve, design operational strategy, 99.5% efficiency curves and a sample simulated trajectory of the wind turbine model. The dashed blue lines are constant wind speed curves and the solid black line indicates beginning of stall.

3. Deriving the Regression Equation

To recap, in below rated operation, and when away from transitional operating states, a wind turbine control system is attempting to track the maximum efficiency curve of the wind turbine [6]. This corresponds to keeping the tip-speed ratio at $\lambda_{\max}$, the tip-speed ratio which corresponds to $C_{\mathrm{p,max}}$. No matter how sophisticated the control system, there will always be some error when attempting to track $C_{\mathrm{p,max}}$, with the true value of λ varying about $\lambda_{\max}$. For example, this can be seen in the sample trajectory of Figure 2. Thus at each time step,

$$\lambda = \lambda_{\max} + \zeta_\lambda, \tag{2}$$

with error term ζ_λ. The aerodynamic torque, Q_{aer}, is given by [3],

$$Q_{aer} = \frac{1}{2}\rho A R v^2 \frac{C_P(\lambda)}{\lambda}. \tag{3}$$

Balancing torque terms gives that the above expression for aerodymanic torque will also be equal to,

$$Q_{aer} = N\hat{Q}_g + J\dot{\hat{\omega}}_r + \mathcal{L}(\hat{\omega}_r), \tag{4}$$

where the losses function $\mathcal{L}$ accounts for drivetrain torque losses; J is rotor inertia, Q_{aer} and Q_g aerodynamic and generator torques respectively, ω_r is generator speed and N the gearbox ratio (the caret symbol is being used to indicate measured values). Manipulation of the above equations then allows us to obtain measured values $\hat{G}$ which it is shown below can be approximated as a polynomial regression equation involving $C_{\mathrm{p,max}}$ and the drivetrain losses;

$$\begin{aligned} \hat{G} \; &:= \; \frac{N\hat{Q}_g + J\dot{\hat{\omega}}_r}{\frac{1}{2}\rho A R^3 \hat{\omega}_r^2} && (5)\\ \text{(from Equations 3 and 4)} \quad &= \; \lambda^{-3}C_P(\lambda) + \mathcal{L}^*(\hat{\omega}_r) && (6)\\ \text{(Taylor exp. about } \lambda = \lambda_{\max}) \quad &\approx \; \lambda_{\max}^{-3}C_{\mathrm{p,max}} + \mathcal{L}^*(\hat{\omega}_r) + \delta, && (7) \end{aligned}$$

where $\mathcal{L}^* = -\mathcal{L}/\frac{1}{2}\rho A R^3 \hat{\omega}_r^2$ and we have assumed ζ_λ is small enough to allow us to approximate the aerodynamic term $\lambda^{-3}C_P(\lambda)$ (via a first order Taylor expansion about $\lambda = \lambda_{\max}$) as being the constant $\lambda_{\max}^{-3}C_{\mathrm{p,max}}$ plus an error term $\delta = m\zeta_\lambda$ with $m = \frac{d}{d\lambda}\left(\lambda^{-3}C_P(\lambda)\right)\Big|_{\lambda=\lambda_{\max}}$. It is further assumed that ζ_λ, and hence δ, is a noise term. Letting Φ denote the constant term $\lambda_{\max}^{-3}C_{\mathrm{p,max}}$, our measured data is then of the form,

$$\hat{G} = \Phi + \mathcal{L}^*(\hat{\omega}_r) + \delta. \tag{8}$$

Assuming torque losses increase linearly with rotational speed (torque losses have been shown to increase with what is essentially linear behaviour in the literature, for example see the loss versus rotational speed diagrams in [5]) it follows that $\mathcal{L}^*$ is quadratic in $\hat{\omega}_r^{-1}$ (with no constant term).

Identification of Φ and $\mathcal{L}^*$ in this case then allows for the losses, $\mathcal{L}$, and the value of $C_{\mathrm{p,max}}$ to be estimated and hence the identification of these terms has been formed into a polynomial regression problem, with all required data coming from information which is available to a wind turbine controller. In order to form the regression measurements $\hat{G}$ for a given wind turbine it follows, from Equation 5, that the required data is: torque demand, rotational speed, gearbox ratio, rotor inertia and rotor radius.

4. Gaussian Processes

Gaussian process machine learning was initially developed in the early nineties after GPs were found to be the natural limit of some types of neural networks as the number of nodes tended towards infinity [7]. Modelling a function as a GP involves building multivariate Gaussian distributions between all possible subsets of function outputs across the functional domain [11]. Assuming a prior mean of zero, this amounts to determining a prior covariance function, k, between any two outputs of the given function. Covariances between a function's outputs are generally modelled as being a function of the input variables for each pair of output points [11]. Thus for a function, $f(x)$ say, being modelled as a zero prior mean GP then the covariance between function values $f(x_1)$ and $f(x_2)$ will be modelled as,

$$\mathrm{cov}(f(x_1), f(x_2)) = k(x_1, x_2), \tag{9}$$

for some covariance function k.

Probably the most commonly used covariance function is the *squared-exponential* covariance function,

$$k(x_1, x_2) = a \exp\left(-\frac{d}{2}(x_1 - x_2)^2\right). \tag{10}$$

Note that the parameters a and d determine the amplitude and lengthscale of the covariance function respectively.

For the current work, however, it has been shown that the regression equation takes the form of a polynomial regression. This can also be performed within a GP framework by using the correct covariance function. A polynomial covariance function can be readily derived by assuming that the polynomial coefficients are zero mean, independent random variables, this results in the covariance function shown in Equation 11 [12], where d is the degree the polynomial and the parameters γ_k are the variances of the corresponding polynomial coefficients.

$$k_P(x_1, x_2) = \sum_{k=0}^{d} \gamma_k x_1^k x_2^k. \tag{11}$$

The γ_k parameters are determined prior to making predictions using the measured regression data in a standard maximum-likelihood optimisation procedure [11], the measured data is assumed to be corrupted by independent Gaussian noise and the variance, ξ, of this noise is determined in this same optimisation procedure.

Once this GP prior has been determined, regression is performed by conditioning the prior on the measurements via the Gaussian conditional distribution [11, 12].

5. Simulation and Data Collection

The above derivations have been investigated using simulated data from the Supergen Exemplar 5MW wind turbine model, full details of which can be found in [6]. Summary data for this model is given in Table 1. Both the power curve and torque-speed diagrams, along with the sample trajectory, in Figures 1 and 2 were produced using this model. Simulations were run over a range of wind conditions with mean wind speeds between 5 and 8 m/s, turbulence intensities of between 5 and 20% and a power law shear exponent of 0.2. The relevant data, i.e. that required to determine $\hat{G}$ in Equation 5, was extracted from the model. The turbine simulations were performed in Simulink software and then GP regression and results analysis was done using Matlab.

Table 1. Data for simulated wind turbine.

Rated power	5 MW
Rotor diameter	126 m
Blade number	3
Hub height	90 m
Aerodynamic control	Pitch
Fixed/Variable speed	Variable
Pitch Controller	PI with low pass filter
Below Rated Torque Control	Closed loop

As can be seen in the sample trajectory of Figure 2, correlations exists for short time-scales in the turbine measurements. Autocorrelation investigations of the measured values, $\hat{G}$, have shown 20s to be a suitable sampling time in order to avoid having correlated measurements for the current model. Hence, data points were extracted at intervals of 20s during simulation.

Polynomial regression was then performed on these measured values using GP machine learning. For the current case we are performing a quadratic polynomial regression, hence the GP covariance function is,

$$k_P(x_1, x_2) = \gamma_2 x_1^2 x_2^2 + \gamma_1 x_1 x_2 + \gamma_0. \tag{12}$$

In order to provide a benchmark, standard least-squares (LS) polynomial regression is also performed using the inbuilt Matlab function *polyfit*. Finally, from the results of both of these regression techniques were extracted the given estimates of the $C_{\mathrm{p,max}}$ value and drivetrain-losses function. This is done by scaling the given coefficients by the appropriate terms, for example, in order to obtain the $C_{\mathrm{p,max}}$ estimate from the constant coefficient Φ, as determined by regression, one simply multiplies by λ^3_{max}.

6. Results

This regression problems was found to require a robust (in the statistical sense) method for its solution. This is due to the characteristics of the measured data which have been found to contain outliers, has heavy tails and the noise term present is strongly non-Gaussian; all factors which make finding accurate solutions more difficult. Figures 3 and 5 show the values of $C_{\mathrm{p,max}}$ and the drivetrain loss functions respectively predicted by both LS and GP regression techniques. Each prediction is from a dataset containing 500 points, corresponding to roughly 3 hours of realtime operation.

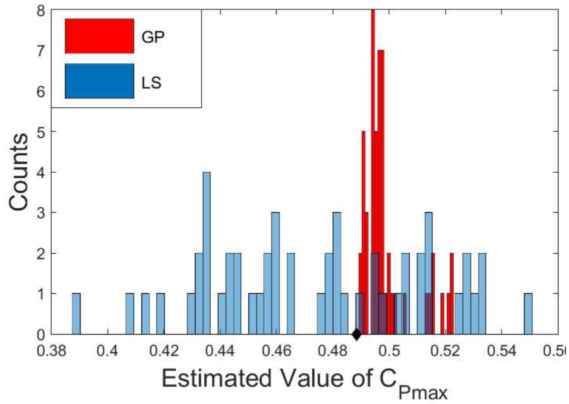

Figure 3. $C_{\mathrm{p,max}}$ estimates from both GP and LS. The true value is given by the black diamond.

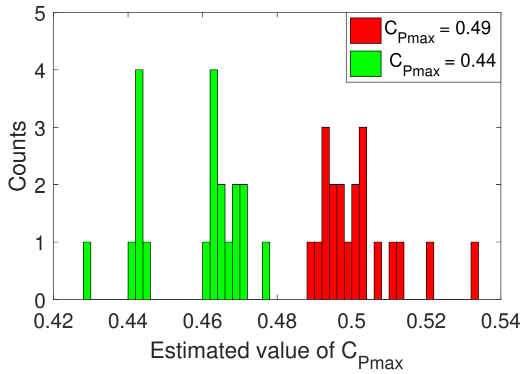

Figure 4. $C_{\mathrm{p,max}}$ estimates from datasets with two different values of maximum efficiency.

In Figure 3 it is clear that the GP approach is superior to using LS. While the GP predictions are clustered close to the true value of $C_{\mathrm{p,max}}$, with some amount of positive bias, the LS results on the other hand are spread almost uniformly across the range of probable values. Based on the GP clustering, one would expect a shift in $C_{\mathrm{p,max}}$ by some small amount to be detectable. In order to test this hypothesis, regression was performed on a dataset generated with all C_{p} values reduced by 0.05 and compared to regression on the original data. $C_{\mathrm{p,max}}$ estimates from these two cases are shown in Figure 4 where the sets of prediction clusters are clearly separate, and hence the GP predictions are indeed able to detect this shift in aerodynamic efficiency. Note that the positive bias in the GP predictions appears to be from the noise distribution being non-symmetrical.

Similar results are then observed for the drivetrain loss predictions in Figure 5 where the true losses in the model are shown along with various GP and LS regression predictions. Again, the GP predictions are much more tightly clustered and, while they do show a bias, GP regression here can be seen to give both more accurate and more consistent predictions of the losses in the drivetrain.

The GP method was then further tested with respect to sensitivity to drivetrain losses by adjusting the losses function in the model. Figure 6 shows the GP predictions from Figure 5, using the depicted true losses function. This set of prediction has been labelled cluster A in Figure 5. Regression was then performed for data generated with an altered loss function

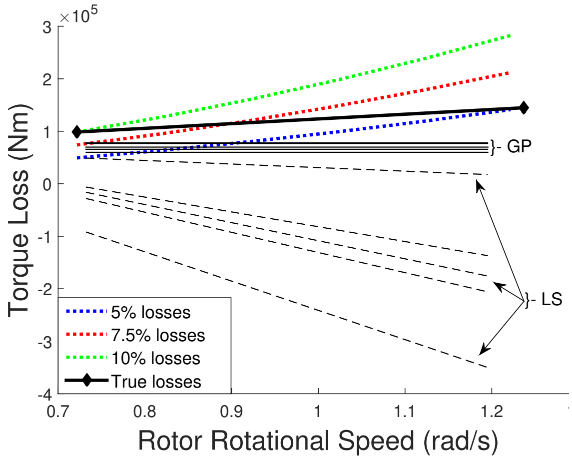

Figure 5. Drivetrain loss predictions along with loss contours (percentages in terms of design power values at each rotational speed).

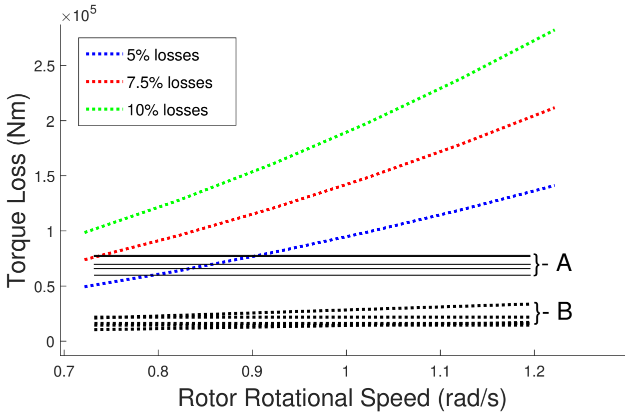

Figure 6. Drivetrain loss predictions along with loss contours for two different loss function cases.

obtained by shifting the left hand value of the original loss function down by 2.5%. This second set of predictions is labelled cluster B. It can be seen that these two clusters are clearly separate and so this shift in losses is detectable when using the GP regression approach.

7. Discussion

The results presented here first demonstrate that GP is superior to LS in this case, with LS results being so scattered as to make them effectively useless. The GP results, while more tightly clustered, do suffer from a bias which appears to be due to assymetrical noise being present. However, as discussed in Section 1, the identification of a turbine's maximum efficiency coefficient and drivetrain losses in below rated operation is being developed primarily for turbine monitoring and O&M purposes; for this type of application it is the ability to detect *changes* in these values, rather than the exact values themselves, which is key to forewarning of potential issues. Hence, for the desired application these offsets do not pose a significant problem. Nevertheless, in future work attempts to remove these biases will be made, focussing on the λ terms which will vary asymmetrically with wind speed and hence may be a strong source of bias here. If this is found to be the case then it may be possible to reformulate the regression problem in order to account for the bias term.

When considering GP predictions of both $C_{\mathrm{p,max}}$ and drivetrain losses, the current work has demonstrated that GP regression is sensitive enough to detect changes in these values. The next stage in developing this method will therefore be to devise an automated strategy for determining when perceived changes in predictions are deemed to be significant. This will most likely be done by implementing the probabilistic aspects of GP models in order to give confidence intervals for the predictions and so probabilistic thresholds can be used to indicate when a likely shift in value has been detected.

8. Conclusions

The identification of a wind turbine's $C_{\mathrm{p,max}}$ value and drivetrain-losses was formulated as a polynomial regression problem for which the relevant data is available to a wind turbine controller. GP machine learning was shown to give superior performance in this problem as compared to LS. The GP results indicate that the approach presented here could prove to be a

IOP Conf. Series: Journal of Physics: Conf. Series **1037** (2018) 032024 doi:10.1088/1742-6596/1037/3/032024

useful and inexpensive tool for detecting changes in wind turbine dynamics for monitoring and O&M purposes. This method is also attractive because it does not require any new sensors to be installed, since all required data is already available to the wind turbine controller. Future work on this method will focus on removing the prediction biases and validating the results seen here using real turbine data.

References

[1] Wang X, Bibeau E, and Naterer G 2007 *Experimental investigation of energy losses due to icing of a wind turbine*, International Conference on Power Engineering, Hangzhou, China

[2] Sareen A, Sapre C and Selig M 2014 *Effects of leading edge erosion on wind turbine blade performance*, Wind Energ. 17:15311542

[3] Burton T, Jenkins N, Sharpe D and Bossanyi E 2011 *The Wind Energy Handbook*, Second Edition, John Wiley & Sons

[4] Bianchi F, Battista H and Mantz R 2007 *Wind Turbine Control Systems*, Advances in Industrial Control, Springer-Verlag London

[5] Martins R, Fernandes C, Seabra J 2015 *Evaluation of bearing, gears and gearboxes performance with different wind turbine gear oils*, Friction 3(4): 275286

[6] Stock A 2016, *Augmented control for flexible operation of wind turbines*, PhD Thesis, EEE, University of Strathclyde, Glasgow

[7] Neil R 1994 *Priors for infinte networks*, Technical Report CRG-TR-94-1, University of Toronto, Canada

[8] Chen N, Qian Z, Nabney I and Meng X 2014 *Wind power forecasts using gaussian processes and numerical weather prediction*, IEEE Trans. Pow. Sys. 29-2

[9] Papatheou E, Dervilis N, Maguire E and Worden K 2014 *Wind turbine structural health monitoring: a short investigation based on SCADA data*, 7th European Workshop on Structural Health Monitoring, Jul 2014

[10] Seng K 2008, *Non-Linear Dynamics Identification Using Gaussian Process Prior Models Within a Bayesian Context*, PhD Thesis, Hamilton Institute, National University of Ireland

[11] Rasmussen C E and Williams C 2006, *Gaussian Processes for Machine Learning*, MIT Press

[12] Mackay D 1998, *Introduction to Gaussian processes*, NATO ASI Series F Computer and Systems Sciences, 168, p133-166